

CrowdMI: Multiple Imputation via Crowdsourcing

Lovedeep Gondara

School of Computing Science, Simon Fraser University, Burnaby, BC, Canada
lgondara@sfu.ca

Abstract

Can humans impute missing data with similar proficiency as machines? This is the question we aim to answer in this paper. We present a novel idea of converting observations with missing data to a survey questionnaire, which is presented to crowdworkers for completion. We replicate a multiple imputation framework by having multiple unique crowdworkers complete our questionnaire. Experimental results demonstrate that using our method, it is possible to generate valid imputations for qualitative and quantitative missing data, with results comparable to imputations generated by complex statistical models.

1 Introduction

Missing data is unavoidable and is a significant issue. Data can be missing due to any number of reasons, including but not limited to a faulty apparatus, error prone manual data entry, non response in surveys etc. Irrespective of the cause, missing data is always undesirable, especially during analytics phase. Even small proportions of missing data can seriously bias inference, leading researchers to draw wrong conclusions.

Imputing missing values is the most often used method to deal with missing data, where we replace missing values with most probable ones. Imputation is performed using complex modelling strategies such as methods based on regression, Markov Chain Monte Carlo, Bayesian approaches and machine learning. Instead of imputing a single value per missing observation, it is recommended and preferred to impute multiple slightly different values creating multiple versions of the complete dataset. It is done to induce variability in imputation process that accounts for imputation error [Rubin, 2004; Schafer, 1999]. This process is known as **multiple imputation**. Imputed datasets are then analyzed separately and the results combined using the methods of Little and Rubin [Little and Rubin, 2014]. Combined results reflect the error in imputation process.

Recently crowdsourcing has emerged as an efficient platform for collecting information that is otherwise not available [Franklin *et al.*, 2011]. It is a promising concept, with its suc-

cess attributed to unique human attributes such as intuition and large degrees of freedom.

Inspired from the unique human abilities, we propose a novel multiple imputation model based on crowdsourcing, called CrowdMI. CrowdMI works by structuring missing data as a survey questionnaire, which is then completed by crowdworkers. This study is aimed at answering the following questions:

1. Can humans fill in (impute) missing qualitative and quantitative data?
2. If the same missing value is imputed multiple times by humans, will it have similar variation as in values imputed by machine based multiple imputation methods?
3. How much information is needed by crowdworkers for imputing missing observations?

Our contributions in this study are as following:

1. To best of our knowledge, we present the first study to use crowdsourcing for systematic missing data imputation.
2. We present the novel idea of structuring missing data as a survey questionnaire, making it easier for crowdworkers to impute missing values.
3. We present feasibility analysis of using human computation for multiple imputation.
4. We study the impact of presenting crowdworkers with varying degrees of prior knowledge about the dataset before imputing missing values.
5. We present the comparison of human imputed values with machine imputation.

Our study has far reaching impact, successful imputation of qualitative and quantitative data by humans can open research possibilities in active learning. Where observations with missing data can be first imputed by humans before commencing the learning task. Existing crowdsourcing frameworks can leverage human computation to fill in quantitative and qualitative information in databases.

Next section reviews some related work to missing data and crowdsourcing. Section 3 introduces CrowdMI. Section 4 presents evaluation and several challenges using real life datasets. Finally, we conclude the paper in Section 5.

This study has been approved by Office of research Ethics at XXX (Institutional Review Board Approval Number: XXX).

2 Related work

Most common and easiest to use method to deal with missing data is *complete case analysis* in which any instance with missing data is deleted from final analysis. Filling in plausible values using mean or most frequent label from observed data is also common. There are number of studies in statistics and machine learning [Stephens and Scheet, 2005; Efron, 1994; Yuan, 2010; Jerez *et al.*, 2010; Allison, 2001] that propose advanced methods for missing data imputation.

Very few studies in crowdsourcing have attempted to leverage human computation for missing data. CrowdDB [Franklin *et al.*, 2011] uses crowdsourced queries, which can include searching and filling in missing values such as an address or email for a person. Ye *et al.* [Ye and Wang, 2014] proposed a human-machine hybrid approach, involving a model imputing multiple missing values and an oracle(human) selecting the best fit. This severely limits human degrees of freedom, which in turn will limit the models performance, analyzing all imputations as a machine based multiple imputation model is a better alternative in such scenarios.

3 CrowdMI

In this section we introduce CrowdMI, our multiple imputation model based on human computation. First and the most important challenge we face is to present available raw data to a presumably data naive crowdworker, so the crowdworker can fill in the missing values. It is well understood that presenting rows of raw data without proper context to crowdworkers will not yield good results, instead it will inflate the total time needed to complete the job and job costs. Surveys are known to be one of the best methods to extract human knowledge [De Leeuw, 2005; Wright, 2005]. We chose the same route, a survey questionnaire is designed that presents some preliminary information about the dataset and then each row with missing data is presented as an easily interpretable survey question. In a basic format, our process is presented using Figure 1 and Algorithm 1.

Algorithm 1 CrowdMI: Human powered multiple imputation

Require: n :total missing observations, k :number of imputations

```

1: procedure CROWDMI( $n, k$ )
2:   while  $i < k$  do
3:     while  $n \neq 0$  do
4:       select a row from  $n$ 
5:       convert it into survey question
6:       collect answers as  $a$ 
7:   return  $a$ 
```

We demonstrate the process by example using sample dataset shown in Table 1.

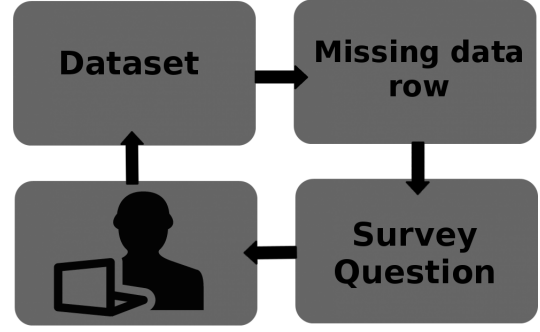


Figure 1: CrowdMI mechanism of action: Missing data rows are extracted from dataset and are converted into a survey question that provides details about the dataset and available information (varying degrees), it is then answered by human participants using a crowdsourcing framework, answers are treated as imputed values and dataset is updated

Table 1: Sample data from Galton’s height dataset, Father and mother column hold the information for height of subjects parents, with gender holding subjects sex information and height has the information for subjects height.

Father	Mother	Gender	Height
78.5	67	M	73.2
75.5	65.5	M	73.5
75	64	F	68
69.5	64.5	F	63.7

We start by providing context around the dataset as an introduction to the survey. In this case, the tuples belong to the famous *Galton’s height data* [Galton, 1886], investigating the relationship between offsprings and their parents height. A sample survey introduction is framed as follows:

”This dataset presents relationships between heights of parents and their adult children. It is seen that there a positive relationship between parents and their offsprings height, in other words taller parents tend to have taller children. It is also observed that on average males are taller than females.”

Additional details such as variable distribution or relationship between different attributes can be added if needed. Then a survey question is formulated as following:

Q: We have a record for which height of a child is missing, given that the gender is male, height of father is 78.5 inch and mother is 67 inch, what do you think is the most probable height for the child?

A: _____

Same question is asked to k participants to complete one iteration of our multiple imputation model for a single row of missing data. Multiple rows of missing data can be combined as part of a single survey, called a *batch* and several *batches* can be combined into a single *job*.

An obvious question comes to mind: ”How many imputations are needed?”, i.e. what value of k should we use?

Our main goal is to draw imputations at random from posterior predictive distribution of missing data, given observed data. There is no rule of thumb for this decision, however statistical simulations have shown [Graham *et al.*, 2007; von Hippel, 2005] that imputations ranging from 5 to 100 can be needed depending on proportion of missing data. However, about 10 imputations are sufficient for good results.

4 Evaluation

This section presents evaluation of CrowdMI on real life datasets under varying conditions.

4.1 Dataset

We start with an introduction to the dataset used for our initial evaluation. We use a publicly available dataset [Tager *et al.*, 1979] that explores the relationship between respiratory function(measured using Forced Expiratory Volume(FEV)) and smoking. FEV is the amount of air an individual can exhale in first second of forceful breath. Data includes measurements on FEV(litres), age(years), height(inches), gender(male/female) and smoke(yes/no). First few rows of dataset are shown in Table 2 for better understanding.

Table 2: FEV data, first column shows the age of the subject, second column holds the information for FEV, third and fourth column show information for height and gender of the subject and last column shows smoking status

Age	FEV	Height	Gender	Smoke
9	1.708	57.0	F	No
8	1.724	67.5	F	No
7	1.720	54.5	F	No
9	1.558	53.0	M	No

4.2 Setup

As it is clear from last section, we are using a dataset with clinical information. The dataset does not have simple and intuitive information and explanation like the height dataset in previous sections. We decided to use this dataset on purpose, we assume that most of our crowdworkers are data naive, or naive enough not to understand the mechanics behind FEV and its relationship with other attributes. This is the perfect setup for our novel study, as if we get good results on this dataset, we can assume better performance on relatively *easy* datasets.

We start our experiment by setting ten observations to have missing values for age and gender at random and pose the imputation problem as a survey question to crowdworkers. We chose age and gender to get an idea of imputation process with different data types. CrowdFlower is used as a crowdsourcing platform in all of our experiments.

4.3 Model calibration

Our first challenge is to calibrate our imputation model, for which we need to understand the type of responses we receive and amount of attention a crowdworker pays to the survey details. For our first job, we created a simple questionnaire with

a basic data description, stated as following:

”In a data set, we have subjects with age from 3 to 19 years, with half over 10 years of age. We also have about 51% males and 49% females. We have a case that has age and gender missing, based on information provided please fill in the values you think are most probable.”

Radio buttons were provided to crowdworkers to select gender and a free text field to enter values for age. Under the free text field, we provided a help text reminding crowdworkers of the valid range for age. The survey was run with 100 *judgements*, where a judgement is defined by CrowdFlower as a valid answer. Each judgement has to be completed by a unique user to prevent users from completing our survey multiple times. We set no constraints over user ability, that is all categories of users from beginner to expert were able to complete this initial survey. We used initial compensation of \$0.1 per judgement. Our expectation was to get a uniform distribution for age within our predefined age limit and gender distribution similar to the one described in initial dataset description.

Initial results were very disappointing, returned age distribution was out of our predefined age limit and gender distribution was biased towards males. Initial results of this survey are shown in Figure 2 and Table 3.

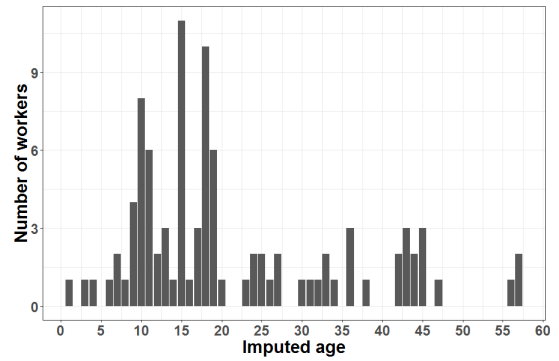


Figure 2: Age distribution from imputation, plot shows crowdworkers did not adhere to guidelines or data description.

Table 3: FEV data, imputed gender distribution, it is evident that crowd imputation was biased towards ”male” for gender

Male	Female
83(83%)	17(17%)

Results show that crowdworkers do not adhere to guidelines or suggestions and try to complete the task as soon as possible. The poor performance can also be attributed to a pool of workers without much experience, having a net negative impact on imputations.

To answer some of the questions, and to measure the impact of constraints on range of crowdworkers input, we reran the same experiment, but this time using a conditional constraint where users were only allowed to enter age from 3 to 19 years, forcing them to stay within the limits of our age

distribution. Results were better compared to the last run, with still the imputation distribution centered around the maximum allowed age value. Imputation proportion for "female" gender increased by 7%, i.e. 17% to 24%. To measure the effect of crowdworkers experience on imputation process, we added an additional constraint where only crowdworkers at highest experience level were allowed to complete the survey. Results were better using most experienced batch of crowdworkers, age distribution was more closer to empirical distribution with 34% less than or equal to 10. Gender distribution improved as well, with 32% imputations for female. So, for the rest of our experiments, we decided to use most experienced crowd with compensation fixed at \$0.15 per judgement and we restrict judgements for each question to be 30, which means 300 total judgements for 10 missing values.

From the results of default survey administered by Crowd-Flower to gather crowd responses for quality and ease of work, and compensation. Our surveys ranged from "easy" to "average" for question difficulty, compensation is deemed "fair" by crowdworkers and instructions for survey completion are reported as "clear". Even though we use most experienced crowd with difficult questions, we had a net "positive" response with over 90% judgements received within first 3 hours.

4.4 Information selection

Our next big challenge is to decide what and how much information should be given to the crowdworkers to facilitate better answers, what criteria should be used to select attributes that are related to the missing data and how should the information be framed as an easy to read survey description. We decided to use a very basic approach of descriptive statistics, where we investigate the relationship with missing attribute and the rest of the attributes in the dataset and provide the information on the attributes that have strongest correlation with missing attribute. This is in principle similar to how regression based imputation works.

Setting ten random rows to have missing age values. We start with the basic description of inter variable relationships, we structure our survey as following:

"This data concerns FEV (Forced Expiratory Volume), a measure of lung functionality in participants 3 to 19 years of age. We know that:

- FEV increases with age and height
- Minimum and maximum FEV in our case is 0.79 and 5.79
- For a 5 year old, average FEV is 1.6 and for a 10 year old, average FEV is 2.7.
- A participant between heights of 55 and 60 inch have average FEV of 2.0 and participants taller than 70 inches have average FEV of 4.3
- Females have slightly lower FEV than males which averages at 2.5 in females compared to 2.8 in males."

We also provided supplemental information in form of two scatter plots, shown in Figure 3.

This is to be noted that we have not provided descriptives that are missing attribute specific, but instead have given a

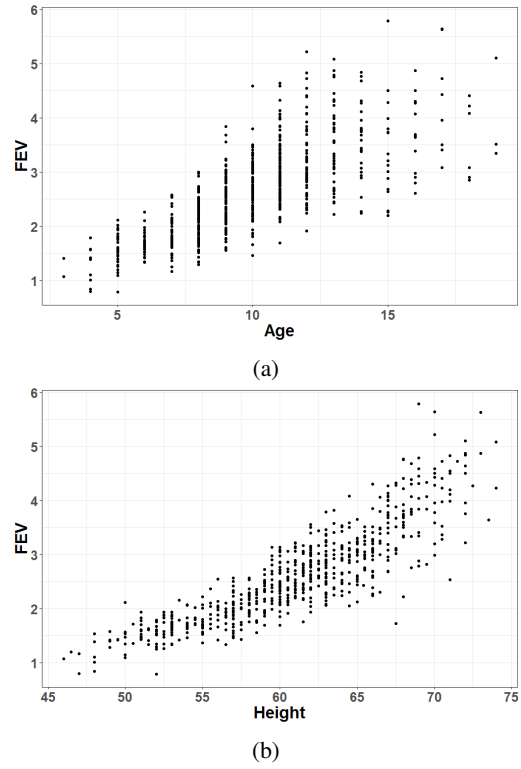


Figure 3: Scatter plots presented to crowdworkers for visual aid to impute missing values.

general overview of the dataset that includes the relation of missing attribute with other variables. This is done on purpose to estimate crowd capacity to reason relationships between different variables, so to generalize this method for cases when more than one variables are missing.

Then we presented questions to crowdworkers that followed format similar to one given below:

Q:"What is the age given that FEV is 1.7 and height is 51, gender is female".

4.5 Comparison with machine imputation

To get an estimate of how good CrowdMI imputations are, we need a competitor. One of the state-of-the-art methods in multiple imputation is Predictive Mean Matching (PMM) [Rubin, 1986; Little, 1988] in Multiple Imputation via Chained Equations (MICE) [Buuren and Groothuis-Oudshoorn, 2011]. It is a complex statistical model and works as following: Given some variable x with missingness and a set of complete variables z , PMM works by regressing x on z producing regression coefficients β . Random draws are then made from posterior predictive distribution of β 's, creating a new set of coefficients β^* . Using β^* , PMM then generates predictive values for x for all cases and for cases where x is missing it chooses a set of cases with observed x whose predicted value for x is closer to predicted values for x with missingness. Then from the chosen k close cases, it randomly selects one and uses it as an imputed value. To get comparable results with CrowdMI,

we ran PMM based imputation to impute 30 datasets as part of multiple imputation.

We chose the state-of-the-art imputation model for one main reason. If crowdworkers can perform at par with PMM for imputation tasks. Then there is an obvious research direction and advantage in missing data scenarios where it is better to use crowdsourcing to gather missing data compared to machine based models. Imputation results from crowd imputed data and PMM with original values are shown in Table 4. As the results are from multiple imputations (30 imputations), results are displayed using median imputed value with 25th and 75th percentile of the distribution of imputed values.

Table 4: Imputation results for missing age, results are shown using median(25th and 75th percentile) of imputations for an observation, results show that CrowdMI imputes values much closer to the original value.

Original	CrowdMI	PMM
5	6.0(5.0,7.0)	6.0(6.0,6.8)
10	11.0(8.0,12.0)	12.0(11.0,12.8)
10	10.5(8.0,12.0)	12.5(11.0,17.3)
11	11.5(9.3,13.0)	8.5(8.0,9.0)
10	12.0(10.0,13.0)	12.0(12.0,15.5)
12	13.0(12.0,14.8)	9.0(8.0,9.0)
11	14.0(12.0,15.0)	11.5(11.0,12.0)
14	14.0(11.5,16.0)	12.0(11.0,14.8)
14	16.0(13.0,16.0)	12.0(11.0,13.0)
16	16.0(13.3,17.0)	12.5(11.0,13.8)

Results show that almost all CrowdMI imputations cover missing data distribution with some median imputed values exactly same as missing observations. CrowdMI imputations are more impressive and even less variable as compared to imputations from PMM, which sometimes did not cover missing value between 25th and 75th percentile of all imputations.

We were interested to investigate if provided plots provided any assistance to crowdworkers, as reading plots is not an easy task and visuals can be distracting from the textual information. So we administered the same survey, this time after removing the plots. Distribution of results was similar but coverage of original missing values was worse compared to results obtained with use of plots. Which shows that an effective visualization of data plays a vital role in a good design of human powered imputation framework.

For next step, we decide to create a scenario little more complicated. This time, we randomly set 10 observations to have missing values for gender. We did not provide any additional information to crowdworkers compared to what we already have. But, we added another plot displaying empirical gender distribution and FEV in the dataset, which is shown as Figure 4

We framed our questions as following with radio buttons provided for input.

Q: "What is the gender given that FEV is 2.4 and height is 62.5, age is 11?"

☐ Male

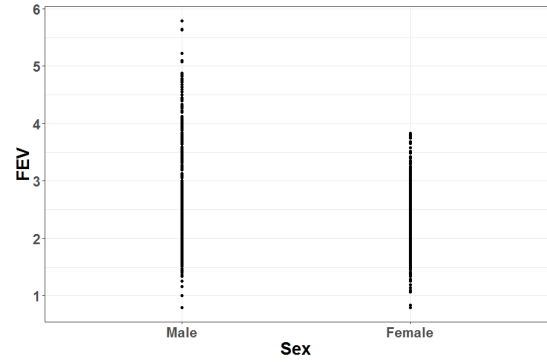


Figure 4: Relation between FEV and gender, additional figure provided to crowdworkers

☐ Female

We would like to state that imputing a categorical variable is more error prone compared to a continuous value, because there is no definition of an error margin or an acceptable range, it is either "True" or "False". We administered the survey for 10 questions with 30 judgements each. To get a comparison with machine imputations, we used PMM with 30 imputations.

Table 5 shows the results. It shows that distribution of imputed values using CrowdMI is very similar to the ones imputed by PMM. Both models agreed on all results but one, where PMM imputed correct value and CrowdMI got it wrong by two votes.

Table 5: Imputation results for missing gender, first column shows original value with second and third showing imputed values by CrowdMI and PMM, values in CrowdMI and PMM show the raw number of votes received by each category from 30 imputations. Both models agree on all observations but one.

Original	CrowdMI(Male - Female)	PMM (Male - Female)
Female	13 - 17	13 - 17
Male	27 - 3	20 - 10
Female	24 - 6	25 - 5
Male	5 - 25	15 - 15
Female	28 - 2	19 - 11
Female	4 - 26	12 - 18
Female	17 - 13	10 - 20
Male	25 - 5	19 - 11
Male	28 - 2	17 - 13
Female	14 - 16	8 - 22

4.6 Perturbed data

For this task, we perturbed some data values from FEV dataset, so the resulting survey is not an exact copy of first survey, but is also not completely different. Running the survey for same number of iterations (30), this time imputation resulted in an overall increased proportion of gender being selected as "male" (about 57% imputed values were male compared to near 40% in last survey), this increase in selected

male proportion can be attributed to perturbation where we decreased age from original value, keeping height the same, which weighs the decision of crowdworkers towards selecting "male". This adds to our claim of good imputation quality where crowdworkers pay careful attention to details.

4.7 Complex dataset

To test CrowdMI on a more complex task, we selected a new dataset, related to diabetes in females of at least 21 years old of Pima Indian heritage [Smith *et al.*, 1988]. Compared to earlier datasets, this dataset is richer in information. It has information on number of times a subject was pregnant, Plasma glucose concentration, Diastolic blood pressure, Triceps skin fold thickness, 2-Hour serum insulin, Body mass index, Diabetes pedigree function, Age and an indicator for diabetic status. Similar to last runs, we randomly set ten observations to have missing values for diabetic status and frame the survey questionnaire as following:

"This data has diabetes information for females aged 21 years and older. We know that:

- Diabetes positive patients have higher blood glucose levels.
- Glucose levels are at average of around 100 for diabetes negative and 130 for positive.
- Diabetes positive also have slightly higher blood pressure, which is about 70 compared to 65 for diabetes negative and
- Blood pressure is also higher for people with higher body mass index.
- Blood pressure is around 77 for people with mass greater than 40 and around 70 for mass less than 40.
- Diabetes positive are also older age compared to negatives, i.e. average age for diabetic is 35 compared to about 25 for non diabetic.
- Diabetics also have on average lower insulin compared to negatives, that is 68 compared to 100 in negatives."

And we also provided three box plots to show the relationship between diabetes status and age, blood pressure, and blood glucose. Factors that can be easily understood by naive audience.

Table 6 shows the results, CrowdMI and PMM disagree on two results, where PMM imputed the correct value. We believe this is due to the fact that PMM is able to use all available data whereas CrowdMI only uses what we decided to share with crowdworkers. Giving more information should improve the outcomes.

4.8 Is human mind Bayesian?

In this section we try to evaluate if given extra information not contained in the dataset, can we influence crowd decision? This is similar to Bayesian methods where we use prior knowledge to update posterior probabilities. In order to do this, for first dataset (FEV), we added a little blurb at end of our questionnaire:

"However, in addition to the information contained in our dataset, we also know that in general population related to this study, females account for about 65% of total."

Table 6: Imputation results for diabetic status, first column shows original values in the dataset, second column is CrowdMI results and third column is results from PMM, Pos is positive, Neg is negative, values in CrowdMI and PMM show the raw number of votes received by each category from 30 imputations. Imputation distribution is similar using CrowdMI and PMM, they disagree on two results.

Original	CrowdMI(Pos - Neg)	PMM(Pos - Neg)
Pos	17 - 13	22 - 8
Pos	25 - 5	21 - 9
Neg	5 - 25	6 - 24
Pos	22 - 8	22 - 8
Neg	5 - 25	2 - 28
Neg	4 - 26	4 - 26
Pos	21 - 9	25 - 5
Pos	5 - 25	18 - 12
Neg	21 - 9	12 - 18
Pos	8 - 22	11 - 19

Results were not significantly different from previous observations, but there was a noticeable shift, with reduced male imputation proportion. We do think that crowdworkers can be further influenced by placing the extra information at start of the questionnaire, as most crowdworkers might not pay full attention to the text below the rest of description.

5 Conclusions, limitations and future work

We conclude our study by answering the questions asked in the introduction:

1. Can humans fill in (impute) missing qualitative and quantitative data?
Our results show that indeed, humans can be used to fill in missing qualitative and quantitative data effectively.
2. If the same missing value is imputed multiple times by humans, will it have similar variation as in values imputed by machine based multiple imputation methods?
Results show that variation induced by multiple imputation using CrowdMI is similar to the variation induced by machine based imputations.
3. How much information is needed by crowdworkers for imputing missing observations?
We have shown that providing a dataset context with some descriptives about attribute correlations and plots for visualization provides an optimal amount of information needed by crowdworkers to fill in missing values.

Limitations and further work: We have only used three datasets (two independent datasets, one perturbed version), in future we would like to validate our method on more datasets. We only considered relatively low dimensional datasets, our future attempts will focus on large dimensional datasets with information presented using a small subset of variables strongly associated with the attribute having missing values.

References

- [Allison, 2001] Paul D Allison. *Missing data*, volume 136. Sage publications, 2001.
- [Buuren and Groothuis-Oudshoorn, 2011] Stef Buuren and Karin Groothuis-Oudshoorn. mice: Multivariate imputation by chained equations in r. *Journal of statistical software*, 45(3), 2011.
- [De Leeuw, 2005] D De Leeuw. To mix or not to mix data collection modes in surveys. *Journal of official statistics*, 21(2):233, 2005.
- [Efron, 1994] Bradley Efron. Missing data, imputation, and the bootstrap. *Journal of the American Statistical Association*, 89(426):463–475, 1994.
- [Franklin *et al.*, 2011] Michael J Franklin, Donald Kossmann, Tim Kraska, Sukriti Ramesh, and Reynold Xin. Crowddb: answering queries with crowdsourcing. In *Proceedings of the 2011 ACM SIGMOD International Conference on Management of data*, pages 61–72. ACM, 2011.
- [Galton, 1886] Francis Galton. Regression towards mediocrity in hereditary stature. *The Journal of the Anthropological Institute of Great Britain and Ireland*, 15:246–263, 1886.
- [Graham *et al.*, 2007] John W Graham, Allison E Olchowski, and Tamika D Gilreath. How many imputations are really needed? some practical clarifications of multiple imputation theory. *Prevention Science*, 8(3):206–213, 2007.
- [Jerez *et al.*, 2010] José M Jerez, Ignacio Molina, Pedro J García-Laencina, Emilio Alba, Nuria Ribelles, Miguel Martín, and Leonardo Franco. Missing data imputation using statistical and machine learning methods in a real breast cancer problem. *Artificial intelligence in medicine*, 50(2):105–115, 2010.
- [Little and Rubin, 2014] Roderick JA Little and Donald B Rubin. *Statistical analysis with missing data*. John Wiley & Sons, 2014.
- [Little, 1988] Roderick JA Little. Missing-data adjustments in large surveys. *Journal of Business & Economic Statistics*, 6(3):287–296, 1988.
- [Rubin, 1986] Donald B Rubin. Statistical matching using file concatenation with adjusted weights and multiple imputations. *Journal of Business & Economic Statistics*, 4(1):87–94, 1986.
- [Rubin, 2004] Donald B Rubin. *Multiple imputation for non-response in surveys*, volume 81. John Wiley & Sons, 2004.
- [Schafer, 1999] Joseph L Schafer. Multiple imputation: a primer. *Statistical methods in medical research*, 8(1):3–15, 1999.
- [Smith *et al.*, 1988] Jack W Smith, JE Everhart, WC Dickson, WC Knowler, and RS Johannes. Using the adap learning algorithm to forecast the onset of diabetes mellitus. In *Proceedings of the Annual Symposium on Computer Application in Medical Care*, page 261. American Medical Informatics Association, 1988.
- [Stephens and Scheet, 2005] Matthew Stephens and Paul Scheet. Accounting for decay of linkage disequilibrium in haplotype inference and missing-data imputation. *The American Journal of Human Genetics*, 76(3):449–462, 2005.
- [Tager *et al.*, 1979] Ira B Tager, Scott T Weiss, Bernard Rosner, and Frank E Speizer. Effect of parental cigarette smoking on the pulmonary function of children. *American Journal of Epidemiology*, 110(1):15–26, 1979.
- [von Hippel, 2005] Paul T von Hippel. Teacher’s corner: how many imputations are needed? a comment on herberger and fisher (2003). *Structural Equation Modeling*, 12(2):334–335, 2005.
- [Wright, 2005] Kevin B Wright. Researching internet-based populations: Advantages and disadvantages of online survey research, online questionnaire authoring software packages, and web survey services. *Journal of Computer-Mediated Communication*, 10(3):00–00, 2005.
- [Ye and Wang, 2014] Chen Ye and Hongzhi Wang. Capture missing values based on crowdsourcing. In *International Conference on Wireless Algorithms, Systems, and Applications*, pages 783–792. Springer, 2014.
- [Yuan, 2010] Yang C Yuan. Multiple imputation for missing data: Concepts and new development (version 9.0). *SAS Institute Inc, Rockville, MD*, 49, 2010.